MULTI-SCALE SPACE VEHICLE COMPONENT IDENTIFICATION

William W. L. Lam, Clement C. C. Pang and Nelson H. C. Yung

Department of Electrical and Electronic Engineering
The University of Hong Kong
Pokfulam, Hong Kong SAR, China

ABSTRACT

Vision based vehicle recognition systems have an important role in traffic surveillance. Most of these systems however fail to distinguish vehicles with similar dimensions due to the lack of other details. This paper presents a new scale space method for identifying components of moving vehicles to enable recognition eventually. In the proposed method, vehicles are first divided into multi-scale regions based on the center of gravity of the foreground vehicle mask. It utilizes both the texture scale space and the intensity scale space to determine regions that are homogenous in texture and intensity, from which vehicle components are identified based on the relations between these regions. This method was tested on over a hundred outdoor traffic images and the results are very promising.

1. INTRODUCTION

In visual traffic surveillance, most vehicle recognition systems simply distinguish the size, dimension, shape or contour of the vehicle, from which vehicles can be categorized by type [3, 4, 6, 7]. Gupte et al. [3] classified vehicles based on a non-rigid vehicle model, while Lai et al. [4] classified vehicles based on dimension. Tan et al. [6] extracted vehicle by matching model discrimination, and Wei et al. [7] classified vehicles using a parameterized model and neural networks. With improvement from [4], Fung et al. [1] additionally considered vehicle shape for classification based on vehicle motion.

Broadly, these techniques addressed the problem of classifying vehicles of the same size or dimension. However, if different vehicles of similar dimension are present, these methods become inadequate, as similar sized vehicles will most likely be recognized as one class because of the non-utilization of vehicle details. Recently high-tech surveillance cameras are emerging and providing high resolution images. This motivates us to

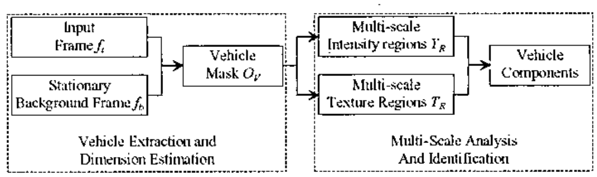


Fig. 1: Scale space identification of vehicle components

consider decomposing the vehicle object into components such as roof, bonnet, wheels and etc, as an alternative way to discriminate between different vehicles. In this paper, we propose a scale space method as depicted in Fig. 1 for this purpose. Assuming that the vehicle is in motion, we first employ the extraction method as described in [5]. From each extracted vehicle, we then estimate its dimension to identify similar sized vehicles and compute its texture scale space (TSS) and intensity scale space (ISS) based on the centers of gravity. By fusing the texture and intensity regions corresponding to a selected scale, vehicle components can be subsequently identified. Details of the proposed method are presented in Section 2, and simulation results and discussion are given in Section 3, where conclusions are drawn in Section 4.

2. PROPOSED METHOD

It is further assumed that the surveillance camera is road-side mounted and stationary, while the light source can be single and strong (as in day time) or multiple and diffused (as in night time). Based on these assumptions, the proposed method comprises of: (1) vehicle extraction and dimension estimation; (2) multi-scale analysis and identification.

2.1. Vehicle Extraction and Dimension Estimation

In vehicle extraction, we first extract the moving vehicles from the stationary background using a texture-based method described in [5]. This method considers the textural difference between the vehicle and other road features and its merits are that it is accurate in generating a vehicle mask (outline describing the vehicle) and able to

0-7803-8554-3/04/$20.00 ©2004 IEEE.

eliminate shadow and reflection on the vehicle chassis. Then, vehicle dimensions are estimated using the deformable model approach as described in [4], which relies on a set of calibrated camera parameters [2]. This set of calibrated camera parameters is useful in providing the transformation between the 3D world coordinates and the 2D image coordinates as seen by the camera. Vehicle dimensions including width, length and height are subsequently calculated in 3D coordinates. This estimation method enables us to identify vehicles of similar dimensions for identification in the next step.

2.2. Multi-Scale Analysis and Identification

After the vehicles have been extracted, the vehicle mask O_V of each similar sized vehicle is constructed by segmenting the boundary of the vehicle through the subtraction between the input frame f_i and the stationary background frame f_b. O_V is symbolized in Fig. 2(a) by the white ellipse. Fig. 2(b-d), illustrate the concept of scale space by partitioning O_V into quarters each times (as in quad-tree methods), using the center-of-gravity (CG) algorithm. This algorithm computes the CG of each region (or sub-region) and divides the region horizontally and vertically into four parts centered at the CG. If O_V is defined as root region R_0 as shown in Fig. 2(a), then this root region divides into four scale 1 sub-regions $R_{0,0}$, $R_{0,1}$, $R_{0,2}$, $R_{0,3}$ as shown in Fig. 2(b) and their corresponding pixels within the vehicle mask O_V are determined as,

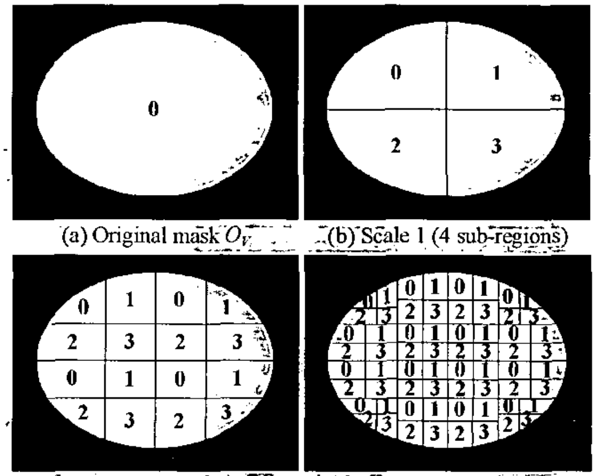
$$x_{0,\ldots,n} = \frac{\sum_{i,j\in R_{0,\ldots,n}} i\cdot O_V(i,j)}{\sum_{i,j\in R_{0,\ldots,n}} O_V(i,j)} \quad y_{0,\ldots,n} = \frac{\sum_{i,j\in R_{0,\ldots,n}} j\cdot O_V(i,j)}{\sum_{i,j\in R_{0,\ldots,n}} O_V(i,j)}$$

$$\begin{aligned} &R_{0,\ldots,n,0}(i,j) \;\; \forall \;\; i \le x_{0,\ldots,n}, j \le y_{0,\ldots,n} \text{ and } i,j \in R_{0,\ldots,n} \\ &R_{0,\ldots,n,1}(i,j) \;\; \forall \;\; i \le x_{0,\ldots,n}, j > y_{0,\ldots,n} \text{ and } i,j \in R_{0,\ldots,n} \\ &R_{0,\ldots,n,2}(i,j) \;\; \forall \;\; i > x_{0,\ldots,n}, j \le y_{0,\ldots,n} \text{ and } i,j \in R_{0,\ldots,n} \\ &R_{0,\ldots,n,3}(i,j) \;\; \forall \;\; i > x_{0,\ldots,n}, j > y_{0,\ldots,n} \text{ and } i,j \in R_{0,\ldots,n} \end{aligned} \quad (1)$$

where n is the sub-region location, i, j are the position displacements and $x_{0,\ldots,n}$, $y_{0,\ldots,n}$ are the central pixel of region $R_{0,\ldots,n}$. The consequent sub-regions $R_{0,\ldots,n}$ can be further partitioned as depicted in Fig. 2(c)-(d). These square like multi-scale regions are useful in comparing visual properties between vehicle components of inconsistent sizes, i.e., the bonnet of the same vehicle that appear to be of different size as seen in different frames within an image sequence.

After partitioning O_V into multi-scale regions, the intensity courier Y_R of each sub-region $R_{0,\ldots,n}$ is determined based on intensity averaging,

$$Y_{R_{0,\ldots,n}} = \frac{\sum_{i,j\in R_{0,\ldots,n}} f_i(i,j)}{\sum_{i,j\in R_{0,\ldots,n}} O_V(i,j)} \quad (2)$$



(a) Original mask O_V (b) Scale 1 (4 sub-regions)

(c) Scale 2 (16 sub-regions) (d) Scale 3 (64 sub-regions)

Fig. 2: Illustration of the partitioning of a mask by the center-of-gravity algorithm

and the textual courier T_R of each sub-region $R_{0,\ldots,n}$ is calculated based on intensity variance,

$$T_{R_{0,\ldots,n}} = \frac{\sum_{i,j\in R_{0,\ldots,n}} \left(f_i(i,j) - Y_{R_{0,\ldots,n}}\right)^2}{\sum_{i,j\in R_{0,\ldots,n}} O_V(i,j)} \quad . \quad (3)$$

Comparing to traditional edge detector of segmentation, our intensity Y_R and textural T_R couriers are helpful in sorting out vehicle components as each vehicle component has comparative intensity or textural relationship with others. From all the T_R at each scale, the mean T_S and moving average $\overline{T_S}$ are calculated by:

$$T_S = \frac{\sum_{n\in S} T_{R_{0,\ldots,n}}}{4^S} \quad \overline{T_S} = \frac{1}{S+1}\sum_{k=0}^{S} T_k \quad , \quad (4)$$

where S is the scale. At texture scale zero, T_S is the largest and continues to decrease in subsequence scales until T_S becomes zero for the smallest sub-region $R_{0,\ldots,n}$. The mean Y_S of all the Y_R at each scale is also calculated by:

$$Y_S = \frac{\sum_{n\in S} Y_{R_{0,\ldots,n}}}{4^S} \quad . \quad (5)$$

To find a suitable scale for identification, we define the critical textural scale C_S, which is chosen when the difference between $\overline{T_S}$ and T_S is maximum,

$$S = C_S \quad \text{if} \quad \max\left(\overline{T_S} - T_S\right) \quad . \quad (6)$$

As depicted in Fig. 3, max($\overline{T_S}$ - T_S) corresponds to a scale (6 in this case) that other scales below or above it gives a smaller variance. If scale lower than C_S is chosen, the component characteristics are insufficient to warrant an accurate identification. On the other hand, if scale higher than C_S is chosen instead, too many superfluous

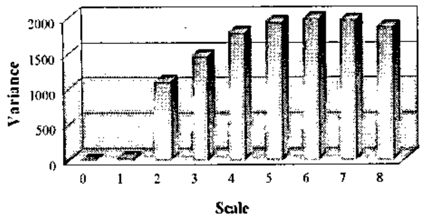
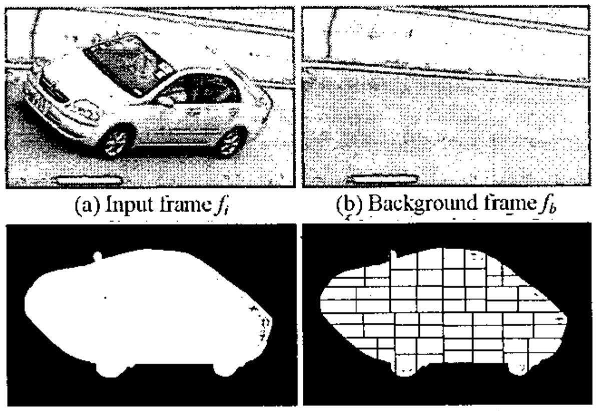


Fig. 3: $\overline{T_S} - T_S$ on each scale of the sedan shown in Fig. 4(a)



(a) Input frame f_i (b) Background frame f_b

(c) Vehicle mask O_V (d) Scale 3 regions

Fig. 4: Partitions on a gray color sedan

features have to be computed and discriminated. Once C_S is determined, vehicle components can be segmented by comparing the related intensity Y_R and textural T_R couriers on the critical textural scale C_S. Such components include body (roof, bonnet, trunk, door panels, fenders, bumpers), wheels, windows, lights and grille.

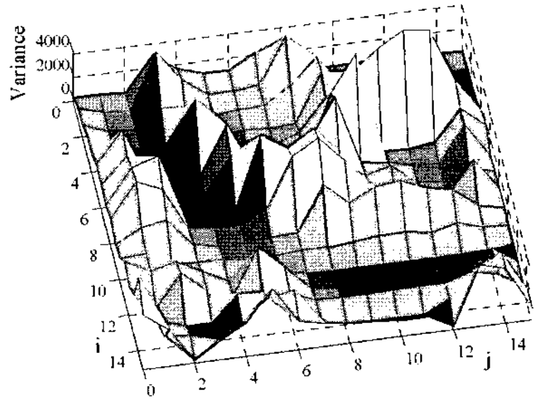
Regions with high textural courier $T_R > T_S$ (edge area) and bounded by low intensity courier $Y_R < Y_S$ (dark area) are defined as wheels, whereas regions with low intensity courier $Y_R < Y_S$ (dark area) and bounded by high textural courier $T_R > T_S$ (edge area) are defined as windows. Region with low textural courier $T_R < T_S$ (flat area) are defined as vehicle body. Finally, regions with high textural courier $T_R > T_S$ (edge area) are defined as lights and grille where light regions contain high intensity courier $Y_R > Y_S$ (bright area), while grille regions contain low intensity courier $Y_R < Y_S$ (dark area).

Once vehicle components are segmented, the geometric or transformed features of these components may be used to describe each component, while the topological relationships between components may also be used as additional feature dimension for identification.

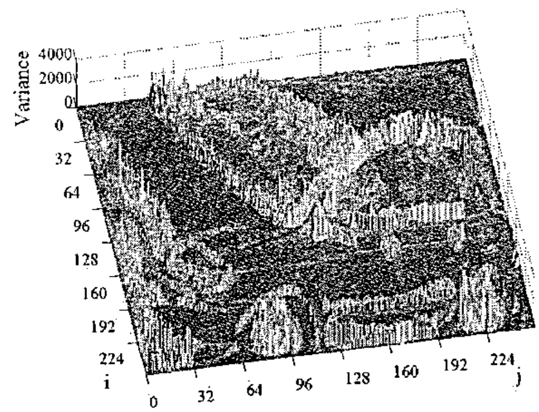
3. RESULTS AND DISCUSSIONS

Over a hundred outdoor traffic image sequences on different roads have been captured under different viewing angles, vehicle sizes and colors for testing the proposed method. Out of all tested images, a gray color sedan sample as depicted in Fig. 4(a) was chosen to illustrate the working of the proposed method. The corresponding background image, vehicle mask and scale 3 regions are depicted in Fig. 4(b), (c) and (d) respectively. With reference to Fig. 3, the critical textural scale C_S occurs at scale 6. In order to show the trend of courier surfaces, Fig. 5(a)-(b) depict the surfaces of texture courier T_R in 3D at scale 4 and scale 8. As can be seen, scale 4 is too coarse to enable a good segmentation, while scale 8 contains too much detail, as discussed in Section 2.2. Fig. 6(a) and (b) depict the textural courier T_R and intensity Y_R courier at scale 6 (C_S), respectively.

From Fig. 6, it can be derived, as illustrated in Fig. 7, those regions with high textural courier and bounded by low intensity courier are wheels. Regions with low intensity courier and bounded by high textural courier are windows. Regions of vehicle body enclose low textural courier. Regions of headlight contain high textural courier

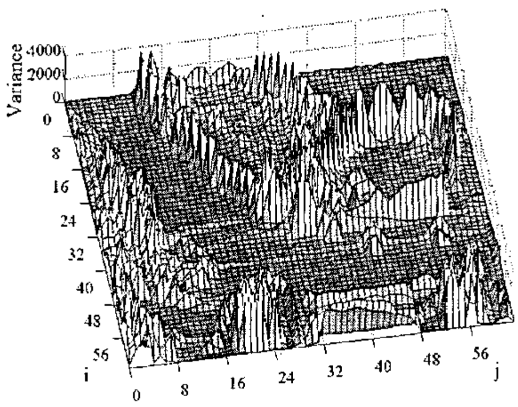


(a) Scale 4

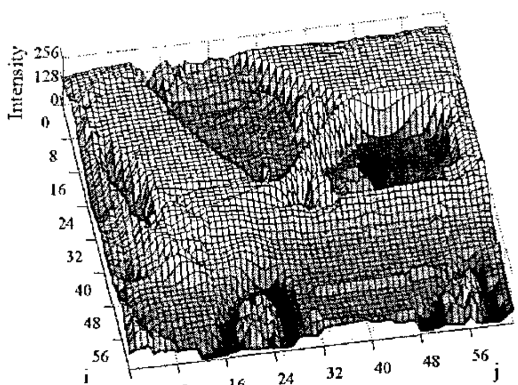


(b) Scale 8

Fig. 5: Texture couriers T_R



(a) Texture couriers T_R at critical texture scale C_S



(b) Intensity couriers Y_R at critical texture scale C_S

Fig. 6: Texture and intensity couriers

and high intensity courier while regions of grille contain high textural courier and low intensity courier. Detailed definitions of each vehicle component are summarized in Table 1. Having this vehicle components representation, further identification of vehicles can be made. For example the colors of two vehicle bodies can be easily compared based on low textural courier at the critical textural scale C_S.

Component	Region Definitions
Wheels	$T_R > T_S$ bounded by $Y_R < Y_S$
Windows	$Y_R < Y_S$ bounded by $T_R > T_S$
Body	$T_R < T_S$
Headlights	$T_R > T_S$ with $Y_R > Y_S$
Grille	$T_R > T_S$ with $Y_R < Y_S$

Table 1: Defined vehicle components based on T_R and Y_R

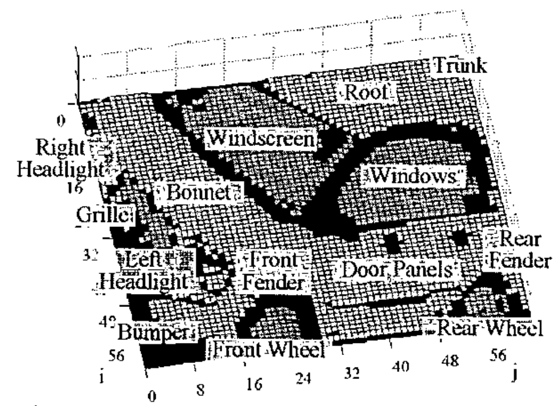


Fig. 7: Vehicle components as described at critical scale C_S

4. CONCLUSIONS

In this paper, we have presented a new scale space method for identifying vehicle components. In this method, textural difference in multi-scale regions is the key to analyzing the vehicle image. Use in conjunction with the intensity difference, vehicle components can be differentiated. In further development, vehicle component details will be utilized for recognition of similar sized vehicles.

5. REFERENCES

[1] G.S.K. Fung, N.H.C. Yung and G.K.H. Pang, "Vehicle shape approximation from motion for visual traffic surveillance," *IEEE Conference on Intelligent Transportation Systems*, pp. 608-613, Aug. 2001.

[2] G.S.K. Fung, N.H.C. Yung, G.K.H. Pang and A.H.S. Lai, "Camera calibration from road land markings," *Optical Engineering*, vol. 42, no. 10, pp. 2967-2977, Oct. 2002.

[3] S. Gupte, O. Masoud and N.P. Papanikolopoulos, "Vision-based vehicle classification," *IEEE Conference on Intelligent Transportation Systems*, pp. 46-51, Oct. 2000.

[4] A.H.S. Lai, G.S.K. Fung and N.H.C. Yung, "Vehicle type classification from visual-based dimension estimation," *IEEE Conference on Intelligent Transportation Systems*, pp. 203-208, Aug. 2001.

[5] W.W.L. Lam, C.C.C. Pang and N.H.C. Yung, "Highly accurate texture-based vehicle segmentation method," *Optical Engineering*, vol. 43, no. 3, pp. 591-603, Mar. 2004.

[6] T.N. Tan and K.D. Baker, "Efficient image gradient based vehicle localization," *IEEE Transactions on Image Processing*, vol. 9, no. 8, Aug. 2000.

[7] W. Wei, Q.S. Zhang and M. Wang, "A method of vehicle classification using models and neural networks," *IEEE Conference on Vehicular Technology*, pp. 3022-3026, May 2001.